# Improving Henderson's method 3 approach when estimating variance components in a two-way mixed linear model

**Razaw Al-Sarraj and Dietrich von Rosen**

**Research Report**
**Centre of Biostochastics**

**Swedish University of Agricultural Sciences**

**Report 2007:8**
**ISSN 1651-8543**

# Improving Henderson's method 3 approach when estimating variance components in a two-way mixed linear model

RAZAW AL-SARRAJ[1] and DIETRICH VON ROSEN

*Centre of Biostochastics*
*Swedish University of Agricultural Sciences*

## Abstract

A two-way linear mixed model, consisting of three variance components, $\sigma_1^2$, $\sigma_2^2$ and $\sigma_e^2$ is considered. The variance component estimators are estimated using a well known non-iterative estimation procedure, Henderson's method 3. For $\sigma_1^2$ we propose two modified estimators. The modification is carried out by perturbing the standard estimator, such that the obtained estimator is expected to perform better in terms of its mean square error. Moreover, Henderson's method 3 can be applied in different ways when decomposing sums of squares. Two different decompositions are considered. The variances of the estimators corresponding to the first random effect are compared to determine which one to choose and to later modify.

**Keywords**: Variance components, Henderson's method 3, perturbed estimators, mean square error, QTL-analysis.

[1]E-mail address to the correspondence author: Razaw.Al-Sarraj@bt.slu.se

# 1 Introduction

In an analysis of variance context, the most commonly used method for estimating the variance components has been through equating the observed and expected mean squares, and solving a set of linear equations. As long as the data are balanced the ANOVA estimators are known to have good statistical properties, i.e., the obtained estimators are unbiased and have minimum variance among all unbiased estimators which are quadratic functions of the observations, see Graybill and Hultquist (1961). However, since real world data often are always unbalanced, this method is no longer appealing. For instance, the uniformly minimum variance property is lost. Furthermore, whether data are balanced or unbalanced, there is nothing in the ANOVA methodology that would prevent negative estimates of the variance components to occur, (LaMotte, 1973).

In a seminal paper Henderson (1953) considered variance component estimation with unbalanced data. He presented three methods of estimation which later on, came to be known as Henderson's method 1, 2 and 3. The obtained estimators are unbiased and translation invariant.

However, since all three methods are variations of the general ANOVA method, they suffer from the weaknesses of it. In particular, the lack of uniqueness.

In this paper we were motivated by Kelly and Mathew's (1994) work, where they improved the ANOVA estimators in a one-way variance component model. The model consists of two variance components, one is the random effect of interest, and the second is the error component. They modified the variance component estimator corresponding to the random effect such that the resulting estimator performed better than the unmodified ANOVA estimator in terms of the mean square error (MSE) criteria. If more components were to be included into the model, they were excluded by orthogonal projections. Hence, the model could always be dealt with as if it had two variance components.

Our aim is to modify the variance component estimators obtained by Henderson's method 3, in a two-way linear mixed model, i.e. a model with three variance components of which two components corresponding to the two random effects included in the model, and the third corresponds to the error component. Here, we want to emphasize that we are primarily interested in one of the variance components. We intend to modify this component and calculate its MSE. Thereafter, we compare it with the MSE of the unmodified one. This modified variance component estimator is expected to perform better in terms of the MSE criteria. Another aim of this paper is to discuss how to apply Henderson's method 3 in practice. With three variance components we will have two natural decompositions. How to choose between them is not clear and will therefore be exploited.

## 1.1 Preparation

Matrices will be used in this work and we need some terminology and notations concerning matrices. A matrix $A$ with $m$ rows and $n$ columns is denoted by $A$: $m \times n$. The element located at the intersection of the $i$:th row and the $j$:th column of $A$ will be denoted $a_{ij}$. A matrix partitioned by its $n$ columns is written as $A = (a_1, a_2, \cdots, a_n)$. Partitioning matrices by rows or other submatrices of proper sizes, are written in the same fashion. The identity matrix is denoted $I$. If the dimension of the identity matrix needs to be emphasized, a lower index will be used, e.g. the $n \times n$ identity matrix is written as $I_n$.

Some important notion used in the subsequent are summarized in the following:

(i) The transpose of $A$ is the matrix $A'$ such that if $A = (a_{ij})$ then $A' = (a_{ji})$.

(ii) A square matrix is symmetric if $A' = A$ holds.

(iii) The trace of a square matrix $A$, $\text{tr}(A)$, is the sum of the diagonal elements of $A$, $\text{tr}(A) = \sum a_{ii}$.

- $\text{tr}(A) = \text{tr}(A')$
- $\text{tr}(AB) = \text{tr}(BA)$
- $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$

(iv) The rank of $A$, is the number of linearly independent columns of $A$.

(v) A square matrix $A$ is positive (negative) definite, abbreviated p.d., (n.d) if for all $x \neq 0$: $x'Ax > 0 \quad (< 0)$.

(vi) An orthogonal matrix $A$ is a square matrix $A$ whose transpose is its inverse: $A'A = AA' = I$.

(vii) A generalized inverse, shortened g-inverse, of $A$: $m \times n$ is any $n \times m$ matrix $A^-$ such that $AA^-A = A$.

(viii) A matrix $A$ is idempotent if $A = A^2$ holds.

- If $A$ is an idempotent matrix, then $\text{rank}(A) = \text{tr}(A)$.
- If $A$ is an idempotent matrix, then the eigenvalues of $A$ consist of ones and zeros.
- $A^-A$ is an idempotent matrix.

(ix) The column space of $A$, denoted by $C(A)$, is the vector space generated by the columns of $A$.

(x) Let $A$ be a real symmetric matrix. Then there exists an orthogonal matrix $\Gamma$ such that $\Gamma' A \Gamma = \Lambda$ or $A = \Gamma \Lambda \Gamma'$, where $\Lambda$ is a diagonal matrix.

We define the mean squared error MSE of an estimator $\widehat{\theta}$, denoted by $\text{MSE}(\widehat{\theta})$, as

$$\text{MSE}(\widehat{\theta}) = \text{D}[\widehat{\theta}] + [\text{Bias}(\widehat{\theta})]^2, \tag{1}$$

where, the variance is denoted by $\text{D}[\cdot]$. The bias of an estimator $\widehat{\theta}$, of a parameter $\theta$ is the difference between the expected value of $\widehat{\theta}$ and $\theta$, i.e., $\text{Bias}(\widehat{\theta})=\text{E}(\widehat{\theta})$-$\theta$.

## 1.2 Quadratic forms

Estimation of variance components for balanced and unbalanced data are based on quadratic forms $Y'AY$ where $A$ is a symmetric matrix, and

$$Y \sim \text{N}(\mu, V).$$

In particular the mean and the variance of $Y'AY$ are needed.

(i) The mean of $Y'AY$ is equal to

$$\text{E}(Y'AY) = \text{tr}(AV) + 2\mu' A\mu, \tag{2}$$

which is true even if $Y$ is not normally distributed.

(ii) The variance of $Y'AY$ is

$$\text{D}[Y'AY] = 2\text{tr}(AVAV) + 4(\mu' AVA\mu). \tag{3}$$

(iii) If $AV$ is idempotent, the distribution of $Y'AY$ is given by

$$Y'AY \sim \chi^2(r_A, \frac{1}{2}\mu' A\mu),$$

where $\chi^2(r_A, \frac{1}{2}\mu' A\mu)$ is non-central chi-square distribution, with degrees of freedom equal to $r_A$, i.e., the rank of $A$, and the non-centrality parameter $\frac{1}{2}\mu' A\mu$.

### 1.3 Important criteria for deriving estimators

Consider the following mixed linear model

$$Y = X\beta + Zu + e, \tag{4}$$

where $Y$ is the $N \times 1$ vector of observations, $X$ is a known $N \times m$ matrix, $\beta$ is an $m \times 1$ vector of unknown fixed effect parameters, and $e$ is an $N \times 1$ vector of random error with mean 0 and dispersion matrix $\sigma_e^2 I_N$. The term $Zu$ given in model (5) is a random term that can be partitioned conformably as

$$Zu = \begin{bmatrix} Z_1 & Z_2 & \dots & Z_r \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \end{bmatrix} = \sum_{i=1}^{r} Z_i u_i.$$

Thus, model (4) can be rewritten as

$$Y = X\beta + \sum_{i=1}^{r} Z_i u_i + e, \tag{5}$$

where $Z_i$ is $N \times n_i$ incidence matrix of known elements, $u_i$ is $n_i \times 1$ vector of random effects, with zero mean value and dispersion matrix $\sigma_i^2 I_{n_i}$, $i = 1, \cdots, r$. Further it is assumed that the $u_i$ and $e$ are uncorrelated random variables. Then from (5), $\mathrm{E}(Y) = X\beta$ and the dispersion matrix $\mathrm{V} = D[Y] = \Sigma_{i=1}^{r} Z_i Z_i' \sigma_i^2 + \sigma_e^2 I_N$. The parameters $\sigma_i^2$ and $\sigma_e^2$ are unknown. Since $Zu$ and $e$ are random effects, they can be combined into one random term. Thus (5) can be rewritten as $Y = X\beta + \Sigma_{i=0}^{r} Z_i u_i$ and the dispersion matrix $\mathrm{V} = \Sigma_{i=0}^{r} Z_i Z_i' \sigma_i^2$, where $u_0 = e$, $\sigma_0^2 = \sigma_e^2$ and $Z_0 = I_N$.

To generalize the idea of estimating a single variance component, we consider estimating a linear function of the variance components, $p_0\sigma_0^2 + p_1\sigma_2^2 + \cdots + p_r\sigma_r^2$, where $p_i$ are known, by a quadratic function $Y'AY$ of the random variable $Y$ in (5). The matrix $A$ should be chosen according to some suitable criteria.

(i) Unbiasedness: If $Y'AY$ is unbiased for $\sum_{i=0}^{r} p_i\sigma_i^2$ for all $\sigma_i^2$, then under the restriction $X'AX = 0$,

$$\mathrm{E}(Y'AY) = \mathrm{tr}(AV) = \sum_{i=0}^{r} \mathrm{tr}(AZ_iZ_i')\sigma_i^2 = \sum_{i=0}^{r} p_i\sigma_i^2. \tag{6}$$

i.e., an unbiased estimator is obtained if $p_i = \mathrm{tr}(AZ_iZ_i')$.

(ii) Translation Invariance: $Y'AY$ is translation invariant if it's value is not affected by any change in the fixed effect parameter for the model. If instead of $\beta$ we consider $\gamma = \beta - \beta_0$ as the unknown parameter, where $\beta_0$ is fixed. Then $Y'AY$ is translation invariant if $Y'AY = (Y - X\gamma)'A(Y - X\gamma)$ for all $\gamma$. Thus $AX = 0$. Since $AX = 0$ always implies $X'AX = 0$, we also have the unbiasedness condition satisfied. However, the reverse is not true i.e., unbiasedness does not imply invariance except when $A$ is n.n.d..

(iii) Minimum Variance: The variance of $Y'AY$ under a normality assumption equals

$$\mathrm{D}[Y'AY] = 2\mathrm{tr}[AVAV] + 4\beta'X'AVAX\beta. \tag{7}$$

Under unbiasedness i.e., $AX = 0$, the variance reduces to

$$\mathrm{D}[Y'AY] = 2\mathrm{tr}[AVAV].$$

The mean squared error, defined in (1), of $Y'AY$ equals

$$\mathrm{MSE}[Y'AY] = \mathrm{D}[Y'AY] + [\mathrm{Bias}(Y'AY)]^2. \tag{8}$$

Using the condition for translation invariance $AX = 0$ and unbiasedness $\mathrm{tr}[AZ_iZ_i'] = p_i$, equation (8) reduces to

$$\mathrm{MSE}[Y'AY] = D[Y'AY] = 2\mathrm{tr}[AVAV].$$

Both (7) and (8), under unbiasedness and invariance reduce to $2\mathrm{tr}(AVAV)$.

### 1.4 ANOVA- based methods of estimation

This method is derived by equating the sums of squares in an analysis of variance table to their expected values. Let $\sigma^2$ be the vector of variance components to be estimated in some model, and let $s$ be a vector of sums of squares. Then taking the expected value

$$E(s) = C\sigma^2, \tag{9}$$

where $C$ is a non-singular matrix, the ANOVA estimator of $\widehat{\sigma}^2$ is based on (9) and is the solution to $s = C\widehat{\sigma}^2$, which equal

$$\widehat{\sigma}^2 = C^{-1}s. \tag{10}$$

The expression in (9) can be extended to include not only sums of squares but also any set of quadratic forms. Let $q = (q_1, q_2, \cdots, q_m)'$ be the $m \times 1$ vector of quadratic forms such that

$$E(q) = A\sigma^2, \tag{11}$$

where $\sigma^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2)'$ is the vector of $k \times 1$ variance components and $A$ being an $m \times k$ matrix of known coefficients. Then, if $m = k$ and $A$ is non-singular, (11) will give $\widehat{\sigma}^2 = A^{-1}q$ as an unbiased estimator of $\sigma^2$, as in (10). In cases when there are more quadratic forms than there are variance components to estimate, the following formula gives an unbiased estimator: $\widehat{\sigma}^2 = (A'A)^{-1}A'q$, (see Searle et al. 1992).

## 2 Henderson's three methods

Henderson (1953) presented in his paper three methods of estimation of variance components, currently known as Henderson's method 1, 2 and 3. This paper is considered to be the landmark work of dealing with the problem of estimation of variance components for unbalanced data. For balanced data, variances are usually estimated using the minimum variance estimators based on the sums of squares, appearing in the analysis of variance table. For unbalanced data the situation is different; it is not always clear which mean squares should be used (see Searle 1971). In our work, we will be concentrating on Henderson's method 3, but we shall briefly review the first two methods as well. These methods are sometimes described as being three different ways of using the general ANOVA-method (Searle 1987). They differ only in the different quadratics (not always sums of squares), used for a vector of any linearly independent quadratic forms of observations. All three methods involve calculations of mean squares, taking their expected values, equating them to the observed ones, and then solving the resulting equations in order to obtain the variance component estimators. Some of the merits of the methods is that they are easy to compute, they require no strong distributional assumptions, and by construction these methods yield unbiased estimators. However, the estimators can fall outside the parameter space, i.e., they can become negative. Moreover, the estimators are not unique, because when there are several random effects, the sums of squares for them can be computed in several ways, i.e, corrected for several combinations of other effects. When data are balanced, all three methods reduce to the usual ANOVA-method. For a review of all three methods, see Searle (1968).

## 2.1 Method 1

The method basically involves calculating uncorrected sums of squares, analogous to those used for the analysis of variance for balanced data. In some cases, they are sums of squares, and in others they are quadratic forms of the data that can be negative. This method is easy to compute but can be used only if it is assumed that except for fixed effect parameter $\beta$ in (5), all other elements in the model are uncorrelated variables with means zero and with variances $\sigma_i^2$. Thus, it can be used only for random models and not for mixed ones, which is one of the shortcomings of the method.

## 2.2 Method 2

The purpose of the method is to correct some of the deficiencies of method 1, and to broaden its use to include more general models such as mixed models which involve estimation of both fixed and random effects. Method 2 involves estimation of the fixed effects by least squares, correcting the data in accordance with these estimates and then using method 1 to estimate the variance components. The method cannot be used on models which include interactions between fixed and random effects.

## 2.3 Method 3

This method can be used on mixed models with or without interactions. Instead of the sums of squares that method 1 and 2 use, method 3 uses reductions in sums of squares due to fitting sub-models of the full model, and then equating the reduced sums of squares to their respective expected values. The outcome will be a set of linear estimation equations, which have to be solved in order to obtain the variance component estimators. The drawback with this method is that sometimes more reduction sum of squares are available than necessary to estimate the variance component estimators (see Searle 1987). In other words, occasionally more than one set of estimating equations for the variance components can be computed for one model. From each set we get different estimators of the variance components. Which set of estimators to prefer is not clear, i.e., the variance component estimators are not unique. We will consider the following two-way mixed model with no interaction,

$$Y = X\beta + Z_1u_1 + Z_2u_2 + e, \quad \text{(full model)} \tag{12}$$

where $\beta$ is the fixed parameter vector and $u_1$, $u_2$ are random effect parameters. For this model there are three variance components to estimate, i.e., the variance of the two random effects denoted by $\sigma_1^2$ and $\sigma_2^2$ respectively, and the third is the error variance component denoted by $\sigma_e^2$. We may obtain several sets of estimation equations. The sub-models which may give estimation equations are,

$$Y = X\beta + e, \tag{13}$$

$$Y = X\beta + Z_1 u_1 + e, \tag{14}$$

$$Y = X\beta + Z_2 u_2 + e. \tag{15}$$

Now we present some special notation for reduction sum of squares which was used by Searle (1971, 1987). Let R(.) denote the reduction sum of squares. The sum of squares used for estimation corresponding to the sub-models (13), (14) and (15) can according to this notation be expressed as, $\mathrm{R}(\beta)$, $\mathrm{R}(\beta, u_1)$ and $\mathrm{R}(\beta, u_2)$, respectively. Another notation which will be needed before we write the possible set of equations is R(./.) which is the reduction sum of squares due to fitting the full model (12) minus that of the sub-model. For (12) two sets of estimation equations may be considered

$$\begin{cases} \mathrm{R}(u_1/\beta) \\ \mathrm{R}(u_2/\beta, u_1) \\ \mathrm{SSE} \end{cases} \qquad \text{or} \qquad \begin{cases} \mathrm{R}(u_2/\beta) \\ \mathrm{R}(u_1/u_2, \beta) \\ \mathrm{SSE} \end{cases}$$

where SSE denotes the residual sum of squares. For the first set of estimation equations we define the following partitioned matrices: $[X]$ , $[X, Z_1]$ and $[X, Z_1, Z_2]$. Each reduction R(./.) can be expressed in the form $Y'AY$ for some symmetric matrix $A$. Define the projection matrix $P_w = w(w'w)^- w'$. Thus $P_w$ is an idempotent matrix, for more properties see Schott (1997). Assuming normality all the reduction sum of squares follow a non-central $\chi^2$ distribution and all these reduction sum of squares are independent of each other and of SSE, see Searle (1987). We shall be using the first set of estimation equation in the first part of the work. In the second part, i.e., in section 4, different reductions in sums of squares will be compared. For the first set of equations we need to define the following projection matrices,

$$P_x = X(X'X)^- X', \tag{16}$$

$$P_{x1} = (X, Z_1)((X, Z_1)'(X, Z_1))^-(X, Z_1)', \tag{17}$$

$$P_{x_{12}} = (X, Z_1, Z_2)((X, Z_1, Z_2)'(X, Z_1, Z_2))^-(X, Z_1, Z_2)'. \tag{18}$$

The reduction sums of squares R(./.) can now be obtained as

$$\mathrm{R}(u_1/\beta) = \mathrm{R}(u_1, \beta) - \mathrm{R}(\beta) = Y'(P_{x_1} - P_x)Y,$$

$$\mathrm{R}(u_2/\beta, u_1) = \mathrm{R}(\beta, u_1, u_2) - \mathrm{R}(\beta, u_1) = Y'(P_{x_{12}} - P_{x_1})Y,$$

and

$$\mathrm{SSE} = Y'(I - P_{x_{12}})Y.$$

To apply the procedure, the expected values of the reduction sums of squares are computed. Thereafter the expected values are to be equated to their observed values and by solving the obtained equations the variance components are obtained. The expression for the expected value given in (2), can be used since the dispersion matrix, denoted by $V$ is $V{=}\sigma_1^2 V_1{+}\sigma_2^2 V_2{+}\sigma_e^2 I$, where $V_1{=}Z_1 Z_1'$ and $V_2{=}Z_2 Z_2'$. The following is obtained

$$\mathrm{E}[\mathrm{R}(u_1/\beta)] = \mathrm{tr}(P_{x_1} - P_x)[\sigma_1^2 V_1 + \sigma_2^2 V_2 + \sigma_e^2 I],$$

$$\mathrm{ER}(u_2/\beta, u_1) = \mathrm{tr}(P_{x_{12}} - P_{x_1})[\sigma_1^2 V_1 + \sigma_2^2 V_2 + \sigma_e^2 I],$$

and

$$\mathrm{E}[\mathrm{SSE}] = \mathrm{tr}(I - P_{x_{12}})[\sigma_1^2 V_1 + \sigma_2^2 V_2 + \sigma_e^2 I].$$

The set of calculated reduction sum of squares may be arranged in a vector. Thereafter by equating these expected values to the observed ones we get

$$\begin{bmatrix} Y'(P_{x_1} - P_x)Y \\ Y'(P_{x_{12}} - P_{x_1})Y \\ Y'(I - P_{x_{12}})Y \end{bmatrix} = J \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \sigma_e^2 \end{bmatrix},$$

where

$$J = \begin{bmatrix} \mathrm{tr}(P_{x_1} - P_x)V_1 & \mathrm{tr}(P_{x_1} - P_x)V_2 & \mathrm{tr}(P_{x_1} - P_x)I \\ \mathrm{tr}(P_{x_{12}} - P_{x_1})V_1 & \mathrm{tr}(P_{x_{12}} - P_{x_1})V_2 & \mathrm{tr}(P_{x_{12}} - P_{x_1})I \\ \mathrm{tr}(I - P_{x_{12}})V_1 & \mathrm{tr}(I - P_{x_{12}})V_2 & \mathrm{tr}(I - P_{x_{12}})I \end{bmatrix}.$$

Thus, the estimators of the variance components are

$$\begin{bmatrix} \widehat{\sigma}_1^2 \\ \widehat{\sigma}_2^2 \\ \widehat{\sigma}_e^2 \end{bmatrix} = J^{-1} \begin{bmatrix} Y'(P_{x_1} - P_x)Y \\ Y'(P_{x_{12}} - P_{x_1})Y \\ Y'(I - P_{x_{12}})Y \end{bmatrix}.$$

However, since $P_{x_1} V_1 = V_1$, $P_{x_{12}} V_2 = V_2$ and $P_{x_{12}} V_1 = V_1$, the $J$ matrix reduces to

$$J = \begin{bmatrix} \mathrm{tr}(P_{x_1} - P_x)V_1 & \mathrm{tr}(P_{x_1} - P_x)V_2 & \mathrm{tr}(P_{x_1} - P_x) \\ 0 & \mathrm{tr}(P_{x_{12}} - P_{x_1})V_2 & \mathrm{tr}(P_{x_{12}} - P_{x_1}) \\ 0 & 0 & \mathrm{tr}(I - P_{x_{12}}) \end{bmatrix}.$$

Let

$$A = (P_{x_1} - P_x), \quad B = (P_{x_{12}} - P_{x_1}), \quad C = (I - P_{x_{12}}),$$
$$a = tr(P_{x_1} - P_x)V_1, \quad b = tr(P_{x_{12}} - P_{x_1})V_2, \quad c = tr(I - P_{x12}),$$
$$d = tr(P_{x_1} - P_x)V_2, \quad e = tr(P_{x_{12}} - P_{x_1}), \quad f = tr(P_{x_1} - P_x)). \quad (19)$$

We note that $A$, $B$ and $C$ are idempotent matrices. Using these notations the estimation equations can be written as

$$\begin{bmatrix} \widehat{\sigma}_1^2 \\ \widehat{\sigma}_2^2 \\ \widehat{\sigma}_e^2 \end{bmatrix} = J^{-1} \begin{bmatrix} Y'AY \\ Y'BY \\ Y'CY \end{bmatrix}. \quad (20)$$

The variance component estimator of $\sigma_1^2$, denoted by $\widehat{\sigma}_{u1}^2$ is:

$$\begin{aligned} \widehat{\sigma}_{u1}^2 &= \frac{\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)\mathrm{tr}(I - P_{x_{12}})Y'(P_{x_1} - P_x)Y}{\mathrm{tr}((P_{x_1} - P_x)V_1)\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)\mathrm{tr}(I - P_{x_{12}})} \\ &\quad - \frac{\mathrm{tr}((P_{x_1} - P_x)V_2)\mathrm{tr}(I - P_{x_{12}})Y'(P_{x_{12}} - P_{x_1})Y}{\mathrm{tr}((P_{x_1} - P_x)V_1)\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)\mathrm{tr}(I - P_{x_{12}})} \\ &\quad + \frac{kY'(I - P_{x_{12}})Y}{\mathrm{tr}((P_{x_1} - P_x)V_1)\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)\mathrm{tr}(I - P_{x_{12}})}, \end{aligned} \quad (21)$$

where $k = \mathrm{tr}((P_{x_1} - P_x)V_2)\mathrm{tr}(P_{x_{12}} - P_{x_1}) - \mathrm{tr}(P_{x_1} - P_x)\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)$. Equation (21) simplifies to

$$\begin{aligned} \widehat{\sigma}_{u1}^2 &= \frac{Y'(P_{x_1} - P_x)Y}{\mathrm{tr}((P_{x_1} - P_x)V_1)} - \frac{\mathrm{tr}((P_{x_1} - P_x)V_2)Y'(P_{x_{12}} - P_{x_1})Y}{\mathrm{tr}((P_{x_1} - P_x)V_1)\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)} \\ &\quad + \frac{kY'(I - P_{x_{12}})Y}{\mathrm{tr}((P_{x_1} - P_x)V_1)\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)\mathrm{tr}(I - P_{x_{12}})}. \end{aligned} \quad (22)$$

Using the previous notations we can write $\widehat{\sigma}_{u1}^2$ as

$$\widehat{\sigma}_{u1}^2 = \frac{Y'AY}{a} - \frac{d(Y'BY)}{ab} + \frac{k(Y'CY)}{abc}, \quad (23)$$

where $A$, $B$, $C$, $b$, $c$ and $e$ are defined as in (19). Despite the fact that in our study we will focus on one of the variance components we also give the estimators of the two other components which may be calculated from (20);

$$\widehat{\sigma}_{u_2}^2 = \frac{\mathrm{tr}(I - P_{x_{12}})Y'(P_{x_{12}} - P_{x_1})Y}{\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)\mathrm{tr}(I - P_{x_{12}})} - \frac{\mathrm{tr}(P_{x_{12}} - P_{x_1})Y'(I - P_{x_{12}})Y}{\mathrm{tr}((P_{x_{12}} - P_{x_1})V_2)\mathrm{tr}(I - P_{x_{12}})},$$

$$\begin{aligned}\widehat{\sigma}_e^2 &= \frac{\mathrm{tr}((P_{x_1}-P_x)V_1)\mathrm{tr}((P_{x_{12}}-P_{x_1})V_2)Y'(I-P_{x_{12}})Y}{\mathrm{tr}((P_{x_1}-P_x)V_1)\mathrm{tr}((P_{x_{12}}-P_{x_1})V_2)\mathrm{tr}(I-P_{x_2})} \\ &= \frac{Y'(I-P_{x_{12}})Y}{\mathrm{tr}(I-P_{x_{12}})}. \qquad (24)\end{aligned}$$

### 2.3.1 Mean Square Error of $\widehat{\sigma}_{u1}^2$

Using (1) the mean square of $\widehat{\sigma}_{u1}^2$ equals its variance since $\widehat{\sigma}_{u1}^2$ is an unbiased estimator,

$$\begin{aligned}MSE(\widehat{\sigma}_{u1}^2) &= D[\widehat{\sigma}_{u1}^2] \\ &= D[\frac{Y'AY}{a} - \frac{d(Y'BY)}{ab} + \frac{k(Y'CY)}{abc}] \\ &= \frac{1}{a^2}D[Y'AY] + \frac{d^2}{a^2b^2}D[Y'BY] + \frac{k^2}{a^2b^2c^2}D[Y'CY] \\ &= \frac{2}{a}\mathrm{tr}[AV]^2 + \frac{2d^2}{a^2b^2}\mathrm{tr}[BV]^2 + \frac{2k^2}{a^2b^2c^2}\mathrm{tr}[CV]^2, \qquad (25)\end{aligned}$$

Moreover since all the involved quadratic forms are uncorrelated, $V = \sigma_1^2V_1 + \sigma_2^2V_2 + \sigma_e^2I$ and the MSE equals

$$\mathrm{D}[\widehat{\sigma}_{u_1}^2] = A_1 + A_2 + A_3, \qquad (26)$$

where

$$\begin{aligned}A_1 &= \frac{2}{a^2}[\mathrm{tr}(AV_1AV_1)\sigma_1^4 + 2\mathrm{tr}(AV_1AV_2)\sigma_1^2\sigma_2^2 \\ &\quad + \mathrm{tr}(AV_2AV_2)\sigma_2^4 + 2\mathrm{tr}(AV_1A)\sigma_1^2\sigma_e^2 \\ &\quad + 2\mathrm{tr}(AV_2A)\sigma_2^2\sigma_e^2 + \mathrm{tr}(A^2)\sigma_e^4], \\ A_2 &= \frac{2d^2}{a^2b^2}[\mathrm{tr}(BV_1BV_1)\sigma_1^4 + 2\mathrm{tr}(BV_1BV_2)\sigma_1^2\sigma_2^2 \\ &\quad + \mathrm{tr}(BV_2BV_2)\sigma_2^4 + 2\mathrm{tr}(BV_1B)\sigma_1^2\sigma_e^2 \\ &\quad + 2\mathrm{tr}(BV_2B)\sigma_2^2\sigma_e^2 + \mathrm{tr}(B^2)\sigma_e^4], \\ A_3 &= \frac{2k^2}{a^2b^2c^2}[\mathrm{tr}(CV_1CV_1)\sigma_1^4 + 2\mathrm{tr}(CV_1CV_2)\sigma_1^2\sigma_2^2 \\ &\quad + \mathrm{tr}(CV_2CV_2)\sigma_2^4 + 2\mathrm{tr}(CV_1C)\sigma_1^2\sigma_e^2 \\ &\quad + 2\mathrm{tr}(CV_2C)\sigma_2^2\sigma_e^2 + \mathrm{tr}(C^2)\sigma_e^4].\end{aligned}$$

Thus, the following MSE is obtained:

$$
\begin{aligned}
\mathrm{MSE}(\widehat{\sigma}_{u1}^2) &= [\frac{2}{a^2}\mathrm{tr}(AV_1AV_1) + \frac{2d^2}{a^2b^2}\mathrm{tr}(BV_1BV_1) + \frac{2k^2}{a^2b^2c^2}\mathrm{tr}(CV_1CV_1)]\sigma_1^4 \\
&+ [\frac{2}{a^2}\mathrm{tr}(AV_2AV_2) + \frac{2d^2}{a^2b^2}\mathrm{tr}(BV_2BV_2) + \frac{2k^2}{a^2b^2c^2}\mathrm{tr}(CV_2CV_2)]\sigma_2^4 \\
&+ [\frac{4}{a^2}\mathrm{tr}(AV_1AV_2) + \frac{4d^2}{a^2b^2}\mathrm{tr}(BV_1BV_2) + \frac{4k^2}{a^2b^2c^2}\mathrm{tr}(CV_1CV_2)]\sigma_1^2\sigma_2^2 \\
&+ [\frac{4}{a^2}\mathrm{tr}(A^2V_1) + \frac{4d^2}{a^2b^2}\mathrm{tr}(B^2V_1) + \frac{4k^2}{a^2b^2c^2}\mathrm{tr}(C^2V_1)]\sigma_1^2\sigma_e^2 \\
&+ [\frac{4}{a^2}\mathrm{tr}(A^2V_2) + \frac{4d^2}{a^2b^2}\mathrm{tr}(B^2V_2) + \frac{4k^2}{a^2b^2c^2}\mathrm{tr}(C^2V_2)]\sigma_2^2\sigma_e^2 \\
&+ [\frac{2}{a^2}\mathrm{tr}(A^2) + \frac{2d^2}{a^2b^2}\mathrm{tr}(B^2) + \frac{2k^2}{a^2b^2c^2}\mathrm{tr}(C^2)]\sigma_e^4.
\end{aligned}
$$

Since $\mathrm{tr}(CV_1) = 0$, $\mathrm{tr}(CV_2) = 0$ and $\mathrm{tr}(BV_1) = 0$ . The above can be simplified to

$$
\begin{aligned}
\mathrm{MSE}(\widehat{\sigma}_{u1}^2) &= [\frac{2}{a^2}\mathrm{tr}(AV_1AV_1)]\sigma_1^4 + [\frac{2}{a^2}\mathrm{tr}(AV_2AV_2) + \frac{2d^2}{a^2b^2}\mathrm{tr}(BV_2BV_2)]\sigma_2^4 \\
&+ [\frac{4}{a^2}\mathrm{tr}(AV_1AV_2)]\sigma_1^2\sigma_2^2 + [\frac{4}{a^2}\mathrm{tr}(A^2V_1)]\sigma_1^2\sigma_e^2 \\
&+ [\frac{4}{a^2}\mathrm{tr}(AV_2A) + \frac{4d^2}{a^2b^2}\mathrm{tr}(BV_2B)]\sigma_2^2\sigma_e^2 \\
&+ [\frac{2}{a^2}\mathrm{tr}(A^2) + \frac{2d^2}{a^2b^2}\mathrm{tr}(B^2) + \frac{2k^2}{a^2b^2c^2}\mathrm{tr}(C^2)]\sigma_e^4. \qquad (27)
\end{aligned}
$$

# 3 Perturbing Henderson's equation

In this section, we modify the variance component estimators obtained by Henderson's method 3. This modification is carried out by perturbing the Henderson's estimation equation. Thus, the obtained variance component estimators are biased. Thereafter, by using some suitable criterion, for instance, the MSE, we evaluate the performance of the estimator by comparing it with the MSE of the unmodified estimator. For the estimation equation (20), we define a new class of estimators

$$
\begin{bmatrix} c_1Y'AY \\ c_1d_1Y'BY \\ c_1d_2Y'CY \end{bmatrix} = J^{-1}\begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \sigma_e^2 \end{bmatrix} \qquad (28)
$$

where $J$ is defined in section (2.3), and $c_1 \geq 0$, $d_1$ and $d_2$ are constants to be determined such that it would minimize the leading terms in the MSE of the estimator. The resulting estimator will perform better in terms of MSE since $c_1 = d_1 = d_2 = 1$ gives the same MSE. Thus, the modified variance component estimator of $\sigma_1^2$, denoted by $\widehat{\sigma}_{11}^2$ is

$$\widehat{\sigma}_{11}^2 = \frac{c_1}{a}(Y'AY - \frac{d}{b}d_1 Y'BY + \frac{k}{bc}d_2 Y'CY), \tag{29}$$

where $A$, $B$, $C$, $a$, $b$, $c$ and $d$ are all defined in (19). The MSE of this modified variance component is

$$\mathrm{MSE}[\widehat{\sigma}_{11}^2] = \mathrm{D}[\widehat{\sigma}_{11}^2] + [\mathrm{E}(\widehat{\sigma}_{11}^2) - \sigma_1^2]^2. \tag{30}$$

Since now (20) is perturbed, the estimator is not unbiased, The variance in (29) equals

$$\mathrm{D}[\widehat{\sigma}_{11}^2] = \frac{c_1^2}{a^2}\mathrm{D}[Y'AY] + \frac{d^2c_1^2d_1^2}{a^2b^2}\mathrm{D}[Y'BY] + \frac{k^2c_1^2d_2^2}{a^2b^2c^2}D[Y'CY],$$

since $\mathrm{D}[\widehat{\sigma}_{11}^2]$ has the same structure as (27). Hence the variance of the modified estimator $\widehat{\sigma}_{11}^2$ can be written

$$\begin{aligned}
\mathrm{D}[\widehat{\sigma}_{11}^2] \quad = \quad & [\frac{2c_1^2}{a^2}\mathrm{tr}(AV_1AV_1)]\sigma_1^4 + [\frac{2c_1^2}{a^2}tr(AV_2AV_2) + \frac{2d^2c_1^2d_1^2}{a^2b^2}tr(BV_2BV_2)]\sigma_2^4 \\
& + [\frac{4c_1^2}{a^2}tr(AV_1AV_2)]\sigma_1^2\sigma_2^2 + [\frac{4c_1^2}{a^2}tr(A^2V_1)]\sigma_1^2\sigma_e^2 \\
& + [\frac{4c_1^2}{a^2}tr(AV_2A) + \frac{4d^2c_1^2d_1^2}{a^2b^2}tr(BV_2B)]\sigma_2^2\sigma_e^2 \\
& + [\frac{2c_1^2}{a^2}tr(A^2) + \frac{2d^2c_1^2d_1^2}{a^2b^2}tr(B^2) + \frac{2k^2c_1^2d_2^2}{a^2b^2c^2}tr(C^2)]\sigma_e^4.
\end{aligned} \tag{31}$$

Now we will calculate the bias part of (29), and thus the expectation of $\widehat{\sigma}_{11}^2$ is needed:

$$\begin{aligned}
E[\widehat{\sigma}_{11}^2] \quad = \quad & \frac{c_1}{a}E(Y'AY) - \frac{dc_1}{ab}d_1E(Y'BY) + \frac{c_1kd_2}{abc}E(Y'CY) \\
= \quad & \frac{c_1}{a}tr[A(\sigma_1^2V_1 + \sigma_2^2V_2 + \sigma_e^2I)] - \frac{dc_1d_1}{ab}tr[B(\sigma_1^2V_1 + \sigma_2^2V_2 + \sigma_e^2I)] \\
& + \frac{c_1kd_2}{abc}tr[C(\sigma_1^2V_1 + \sigma_2^2V_2 + \sigma_e^2I)],
\end{aligned}$$

which can be simplified to

$$\begin{aligned}
E[\widehat{\sigma}_{11}^2] \quad = \quad & [\frac{c_1}{a}tr(AV_1) - \frac{dc_1d_1}{ab}tr(BV_1) + \frac{c_1kd_2}{abc}tr(CV_1)]\sigma_1^2 \\
& + [\frac{c_1}{a}tr(AV_2) - \frac{dc_1d_1}{ab}tr(BV_2) + \frac{c_1kd_2}{abc}tr(CV_2)]\sigma_2^2 \\
& + [\frac{c_1}{a}tr(A) - \frac{dc_1d_1}{ab}tr(B) + \frac{c_1kd_2}{abc}tr(C)]\sigma_e^2.
\end{aligned} \tag{32}$$

Thus, the squared bias can be written

$$
\begin{aligned}
(E[\widehat{\sigma}_{11}^2] - \sigma_1^2)^2 &= [(\frac{c_1}{a} tr(AV_1) - 1)\sigma_1^2 + (\frac{c_1}{a} tr(AV_2) - \frac{dc_1d_1}{ab} tr(BV_2))\sigma_2^2 \\
&\quad + (\frac{c_1}{a} tr(A) - \frac{dc_1d_1}{ab} tr(B) + \frac{c_1kd_2}{abc} tr(C))\sigma_e^2]^2. \qquad (33)
\end{aligned}
$$

If we substitute the variance and biased part back into (30), we get the following:

$$
\begin{aligned}
\text{MSE}(\widehat{\sigma}_{11}^2) &= [\frac{2c_1^2}{a^2} tr(AV_1AV_1)]\sigma_1^4 + [\frac{4c_1^2}{a^2} tr(AV_1AV_2)]\sigma_1^2\sigma_2^2 \\
&\quad + [\frac{2c_1^2}{a^2} tr(AV_2AV_2) + \frac{2d^2c_1^2d_1^2}{a^2b^2} tr(BV_2BV_2)]\sigma_2^4 \\
&\quad + [\frac{4c_1^2}{a^2} tr(A^2V_1)]\sigma_1^2\sigma_e^2 \\
&\quad + [\frac{4c_1^2}{a^2} tr(A^2V_2) + \frac{4d^2c_1^2d_1^2}{a^2b^2} tr(B^2V_2)]\sigma_2^2\sigma_e^2 \\
&\quad + [\frac{2c_1^2}{a^2} tr(A^2) + \frac{2d^2c_1^2d_1^2}{a^2b^2} tr(B^2) + \frac{2k^2c_1^2d_2^2}{a^2b^2c^2} tr(C^2)]\sigma_e^4 \\
&\quad + [(\frac{c_1}{a} tr(AV_1) - 1)\sigma_{u1}^2 + (\frac{c_1}{a} tr(AV_2) - \frac{dc_1d_1}{ab} tr(BV_2))\sigma_2^2 \\
&\quad + [\frac{c_1}{a} tr(A) - \frac{dc_1d_1}{ab} tr(B) + \frac{c_1kd_2}{abc} tr(C)\sigma_e^2]^2. \qquad (34)
\end{aligned}
$$

We write the latter expression as below. First let

$$
r = \frac{c_1}{a} tr(AV_2) - \frac{dc_1d_1}{ab} tr(BV_2).
$$

Rewriting it gives the following:

$$
r = \frac{c_1 d}{a} - \frac{dc_1d_1}{a},
$$

where from (19) we have $tr(AV_2) = d$ and $tr(BV_2) = b$. Moreover, let

$$
t = \frac{c_1}{a} tr(A) - \frac{dc_1d_1}{ab} tr(B) + \frac{c_1kd_2}{ab}. \qquad (35)
$$

Hence, the following mean square error is obtained for the modified estimator

$\widehat{\sigma}^2_{u11}$:

$$
\begin{aligned}
\mathrm{MSE}(\widehat{\sigma}^2_{11}) &= [\frac{2c_1^2}{a^2} tr(AV_1AV_1) + (c_1-1)^2]\sigma_1^4 \\
&+ [\frac{4c_1^2}{a^2} tr(AV_1AV_2) + 2(c_1-1)r]\sigma_1^2\sigma_2^2 \\
&+ [\frac{2c_1^2}{a^2} tr(AV_2AV_2) + \frac{2d^2c_1^2d_1^2}{a^2b^2} tr(BV_2BV_2) + r^2]\sigma_2^4 \\
&+ [\frac{4c_1^2}{a^2} tr(A^2V_1) + 2(c_1-1)t]\sigma_1^2\sigma_e^2 \\
&+ [\frac{4c_1^2}{a^2} tr(A^2V_2) + \frac{4d^2c_1^2d_1^2}{a^2b^2} tr(B^2V_2) + 2rt]\sigma_2^2\sigma_e^2 \\
&+ [\frac{2c_1^2}{a^2} tr(A^2) + \frac{2d^2c_1^2d_1^2}{a^2b^2} tr(B^2) \\
&+ \frac{2k^2c_1^2d_2^2}{a^2b^2c^2} tr(C^2) + t^2]\sigma_e^4. \qquad (36)
\end{aligned}
$$

### 3.1 Mean square error comparison

In this section we compare the mean square errors of the modified $\widehat{\sigma}^2_{11}$ and the unmodified estimator $\widehat{\sigma}^2_{u1}$, given by (36) and (27), respectively. We will investigate if $\mathrm{MSE}(\widehat{\sigma}^2_{11}) \leq \mathrm{MSE}(\widehat{\sigma}^2_{u1})$. To do so we compare all coefficients of $\sigma_1^4$, $\sigma_2^4$ and $\sigma_e^4$ and all their cross combinations which appeared in (36) and (27). We will investigate a number of inequalities. If they hold, then the coefficients of the modified estimator $\widehat{\sigma}^2_{11}$ are less than the coefficients of the unmodified one $\widehat{\sigma}^2_{u1}$.

From the terms corresponding to $\sigma_1^4$ in (36) and (27) it follows that we have to investigate if

$$
\frac{2c_1^2}{a^2}\mathrm{tr}(AV_1AV_1) + (c_1-1)^2 \leq \frac{2}{a^2}\mathrm{tr}(AV_1AV_1). \qquad (37)
$$

From the terms corresponding to $\sigma_2^4$ we obtain that

$$
\begin{aligned}
&\frac{2c_1^2}{a^2}\mathrm{tr}(AV_2AV_2) + \frac{2d^2c_1^2d_1^2}{a^2b^2}\mathrm{tr}(BV_2BV_2) + r^2 \\
&\leq \frac{2}{a^2}\mathrm{tr}(AV_2AV_2) + \frac{2d^2}{a^2b^2}\mathrm{tr}(BV_2BV_2), \qquad (38)
\end{aligned}
$$

should be studied, where $r = (\frac{c_1 d}{a} - \frac{dc_1d_1}{a})$ and by assumption $c_1 > 0$. Corresponding to $\sigma_e^4$ we will study the inequality

$$\frac{2c_1^2}{a^2}\mathrm{tr}(A^2) + \frac{2d^2c_1^2d_1^2}{a^2b^2}\mathrm{tr}(B^2) + \frac{2k^2c_1^2d_2^2}{a^2b^2c} + t^2 \leq \frac{2}{a^2}\mathrm{tr}(A^2) + \frac{2d^2}{a^2b^2}\mathrm{tr}(B^2) + \frac{2k^2}{a^2b^2c} \quad (39)$$

where $k = d\mathrm{tr}(B) - b\mathrm{tr}(A)$ and $t$ is defined in (35).

Now the cross combination coefficients of (27) and (36) will be compared. We have first the coefficients of $\sigma_1^2\sigma_2^2$.

$$\frac{4c_1^2}{a^2}\mathrm{tr}(AV_1AV_2) + 2(c_1-1)r \leq \frac{4}{a^2}\mathrm{tr}(AV_1AV_2), \quad (40)$$

where

$$(c_1-1)r = (c_1-1)(\frac{c_1d}{a} - \frac{dc_1d_1}{a}) = \frac{d}{a}(1-d_1)(c_1^2-c_1).$$

Corresponding to $\sigma_1^2\sigma_e^2$ we investigate

$$\frac{4c_1^2}{a^2}\mathrm{tr}(A^2V_1) + 2(c_1-1)t \leq \frac{4}{a^2}tr(A^2V_1), \quad (41)$$

where

$$(c_1-1)t = (c_1-1)(\frac{c_1}{a}\mathrm{tr}(A) - \frac{dc_1d_1}{ab}tr(B) + \frac{c_1kd_2}{ab}),$$

and $A$, defined in (19), is an idempotent matrix. Finally we also study the coefficients corresponding to $\sigma_2^2\sigma_e^2$,

$$\frac{4c_1^2}{a^2}\mathrm{tr}(AV_2) + \frac{4d^2c_1^2d_1^2}{a^2b^2}tr(B^2V_2) + 2rt \leq \frac{4}{a^2}\mathrm{tr}(A^2V_2) + \frac{4d^2}{a^2b^2}tr(B^2V_2), \quad (42)$$

where

$$\begin{aligned} 2rt &= 2(\frac{c_1d}{a} - \frac{dc_1d_1}{a})(\frac{c_1}{a}\mathrm{tr}(A) - \frac{dc_1d_1}{ab}\mathrm{tr}(B) + \frac{c_1kd_2}{ab}) \\ &= \frac{2c_1^2d}{a^2}(1-d_1)(\mathrm{tr}(A) - \frac{dd_1}{b}\mathrm{tr}(B) + \frac{k}{b}d_2). \end{aligned}$$

In order to find appropriate values of $c_1$, $d_1$ and $d_2$ we have chosen to minimize the leading terms in (36), i.e., the terms that involve the coefficients of $\sigma_1^4$, $\sigma_2^4$ and $\sigma_e^4$, respectively. When minimizing the coefficient of $\sigma_1^4$ in (36) the following equation is obtained,

$$\frac{\partial}{\partial c_1}[\frac{2c_1^2}{a^2}tr(AV_1AV_1) + (c_1-1)^2] = 0,$$

with a solution given by

$$c_1 = \frac{1}{\frac{2}{a^2}tr(AV_1AV_1) + 1}. \tag{43}$$

Moreover, minimizing the coefficient of $\sigma_2^4$ gives

$$\frac{\partial}{\partial d_1}[\frac{2c_1^2}{a^2}tr(AV_2AV_2) + \frac{2d^2c_1^2d_1^2}{a^2b^2}tr(BV_2BV_2) + (\frac{c_1 d}{a} - \frac{dc_1d_1}{a})^2] \quad = \quad 0,$$

which implies

$$d_1 = \frac{1}{\frac{2}{b^2}tr(BV_2BV_2) + 1}. \tag{44}$$

Finally, when minimizing the coefficient of the error variance component $\sigma_e^4$ we have to solve

$$\begin{aligned} &\frac{\partial}{\partial d_2}[\frac{c_1^2}{a^2}tr(A^2) + \frac{2d^2c_1^2d_1^2}{a^2b^2}tr(B^2) + \frac{2k^2c_1^2d_2^2}{a^2b^2c} \\ &+\frac{c_1^2}{a^2}(tr(A))^2 - \frac{2dc_1^2d_1}{a^2b}tr(A)tr(B) + \frac{2c_1^2kd_2}{a^2b}tr(A) \\ &-\frac{2dc_1^2kd_1d_2}{a^2b^2}tr(B) + \frac{c_1^2k^2d_2^2}{a^2b^2} + \frac{d^2c_1^2d_1^2}{a^2b^2}(tr(B))^2] \quad = \quad 0. \end{aligned}$$

The minimum is obtained when

$$d_2 = \frac{\frac{d}{b}d_1tr(B) - tr(A)}{(\frac{k}{b})(\frac{2}{c}+1)}. \tag{45}$$

It has been verified that if $c_1$, $d_1$ and $d_2$ satisfy the minimum of the coefficients $\sigma_1^4$, $\sigma_2^4$ and $\sigma_e^4$, respectively, in equation (36). It follows that (37) and (38) hold for the given values in (43) and (44), respectively. Concerning (39), omitting $a^2$ and simplifying, the left hand side can be written as

$$c_1^2tr(A) + \frac{d^2c_1^2d_1^2}{b^2}\mathrm{tr}(B) + \frac{k^2c_1^2d_2^2}{b^2c} + \frac{1}{2}c_1^2\big(\mathrm{tr}(A) - \frac{d}{b}d_1\mathrm{tr}(B) + \frac{kd_2}{b}\big)^2. \tag{46}$$

However, since $c_1$ and $d_1$ given by (43) and (44) respectively, are less than 1 it is enough to study when

$$\frac{k^2c_1^2d_2^2}{b^2c} + \frac{1}{2}c_1^2\big(\mathrm{tr}(A) - \frac{d}{b}d_1\mathrm{tr}(B) + \frac{kd_2}{b}\big)^2 \leq \frac{k^2}{b^2c} \tag{47}$$

The following is obtained after substituting $d_2$ defined in (45) into the left hand side of (47)

$$\frac{k^2c_1^2}{b^2c}\frac{(\frac{d}{b}d_1\mathrm{tr}(B) - \mathrm{tr}(A))^2}{(\frac{k}{b})^2(\frac{2}{c}+1)^2} + \frac{c_1^2}{2}\big(\mathrm{tr}(A) - \frac{d}{b}d_1tr(B) + \frac{k}{b}\frac{\frac{d}{b}d_1tr(B) - \mathrm{tr}(A)}{(\frac{k}{b})(\frac{2}{c}+1)}\big)^2 \tag{48}$$

which can be simplified to,

$$c_1^2(\frac{d}{b}d_1\mathrm{tr}(B)-\mathrm{tr}(A))^2[\frac{c}{(2+c)^2}-\frac{2}{(2+c)^2}]. \tag{49}$$

Hence, for (39) to hold the following must be satisfied

$$(\frac{d}{b}d_1\mathrm{tr}(B)-\mathrm{tr}(A))^2 \leq (\frac{d}{b}\mathrm{tr}(B)-\mathrm{tr}(A))^2. \tag{50}$$

Therefore we have two cases to consider, either

$$\mathrm{tr}(A) \leq \frac{d}{b}d_1\mathrm{tr}(B), \tag{51}$$

or

$$\mathrm{tr}(A) > \frac{d}{b}d_1\mathrm{tr}(B). \tag{52}$$

Which have to be treated separately. If (51) holds, then (50) is always satisfied. If instead (52) is true we will return one step and suppose $d_1 = 1$. Then, obviously (38) and (50) will hold. Observe that $d_1 = 1$ means that we should not perturb (28) with respect to $d_1$.

Moreover, (40) is always satisfied since,

$$(c_1-1)r = \frac{d}{a}(1-d_1)(c_1^2-c_1) \leq 0. \tag{53}$$

Concerning (41), we study the second term in the left hand side,

$$(c_1-1)t = (c_1-1)(\frac{c_1}{a}tr(A) - \frac{dc_1d_1}{ab}tr(B) + \frac{c_1kd_2}{ab}).$$

Substituting $d_2$, defined in (45), yields

$$(c_1-1)(\frac{c_1}{a}tr(A) - \frac{dc_1d_1}{ab}tr(B) + \frac{c_1k}{ab}\frac{\frac{d}{b}d_1tr(B)-tr(A)}{(\frac{k}{b})(\frac{2}{c}+1)}),$$

giving

$$\frac{1}{a}(c_1^2-c_1)\big(tr(A) - \frac{dc_1}{b}d_1tr(B) + \frac{\frac{d}{b}d_1tr(B)-tr(A)}{\frac{2}{c}+1}\big).$$

Thus, for (41), we have from (19) that $\mathrm{tr}(AV_1) = a$ which implies that (41) can be written as

$$\frac{2c_1^2}{a} + \frac{1}{a}(c_1^2-c_1)\big(tr(A) - \frac{dc_1}{b}d_1tr(B) + \frac{\frac{d}{b}d_1tr(B)-tr(A)}{\frac{2}{c}+1}\big) \leq \frac{2}{a}.$$

Hence, if (51) is true (41) will hold if

$$2c_1^2 + (c_1^2 - c_1)(\mathrm{tr}(A) - \frac{d}{b}d_1\mathrm{tr}(B))(\frac{2}{2+c}) \leq 2, \tag{54}$$

and we obtain the additional condition

$$\mathrm{tr}(A) \geq \frac{d}{b}d_1\mathrm{tr}(B) - \frac{(2+c)(1+c_1)}{c_1}. \tag{55}$$

If (52) holds, then it's obvious that (55) will be true. Finally, we check the inequality (42). Since from (19) we have $\mathrm{tr}(AV_2) = d$ and $\mathrm{tr}(BV_2) = b$ we rewrite (42) as

$$\begin{aligned}\frac{4c_1^2 d}{a^2} + \frac{4d^2c_1^2d_1^2}{a^2b^2} \quad &+ \quad \frac{2c_1^2 d}{a^2}(1-d_1)(\mathrm{tr}(A) - \frac{dd_1}{b}\mathrm{tr}(B) + \frac{k}{b}d_2) \\ &\leq \frac{4d}{a^2} + \frac{4d^2}{a^2b}.\end{aligned}$$

It is enough to investigate the third term in the left hand side:

$$\frac{c_1^2 d}{a^2}(1-d_1)(\mathrm{tr}(A) - \frac{dd_1}{b}\mathrm{tr}(B) + \frac{k}{b}d_2).$$

As previously, after substituting $d_2$ and omitting identical terms from both sides, (42) can be written as,

$$2c_1^2 + \frac{2dc_1^2d_1^2}{b} + c_1^2(1-d_1)(\mathrm{tr}(A) - \frac{d}{b}d_1\mathrm{tr}(B))(\frac{2}{2+c}) \leq 2 + \frac{2d}{b}. \tag{56}$$

Thus, (42) is satisfied under (51). Moreover, if $d_1 = 1$ as assumed if (52) holds, then (42) is also valid.

The above results can be summarized in the following proposition

**Proposition 1.** *Let the variance component estimator corresponding to the first random effect* $\widehat{\sigma}_{u1}^2$ *in the model defined in (12) be modified as in (29), where* $c_1$*,* $d_1$ *and* $d_2$ *are chosen as in (43), (44) and (45), respectively. Then (37)– (42) are sufficient conditions for MSE(*$\widehat{\sigma}_{11}^2$*)≤ MSE(*$\widehat{\sigma}_{u1}^2$*).*

Moreover, for the two cases that emerged from (50) we have the following theorem

**Theorem 1.** *Given the model defined in (12), let MSE(*$\widehat{\sigma}_{u1}^2$*) be the mean square error of the unmodified estimator given in (27) and let MSE(*$\widehat{\sigma}_{11}^2$*) be the mean square error of the modified estimator given in (36).*

(i) *If (51) and (55) hold,* $MSE(\widehat{\sigma}^2_{11}) \leq MSE(\widehat{\sigma}^2_{u1})$.

(ii) *If (52) and* $d_1 = 1$, $MSE(\widehat{\sigma}^2_{11}) \leq MSE(\widehat{\sigma}^2_{u1})$.

Note that if Theorem 1 (ii) is applied then the unbiased estimator given in (21) can be modified as in the following:

$$\widehat{\sigma}^2_{11} = \frac{c_1}{a}(Y'AY - \frac{d}{b}Y'BY + \frac{k}{bc}d_2Y'CY). \tag{57}$$

# 4 Variance comparison from two different decompositions of Henderson's method 3

A crucial point when applying Henderson's method 3 is the decomposition of the reduction sums of squares. Unfortunately there is no unique way of how to carry out this decomposition. In this section we compare two choices. In (12), there are three variance components to be estimated. The number of variance components in this model can be reduced, e.g., from three to two variance components by using a suitable transformation method. If an orthogonal vector is defined that is orthogonal to both the fixed effect vector and to, e.g., $Z_2$, so that two variance components $\sigma^2_1$ and $\sigma^2_e$, are estimated instead of three, the estimation problem is simplified, see Khuri et al. (1998). The estimators can thereafter be modified as in Kelly and Mathew (1994). In our work we want to estimate the variance components for (12), by dealing with the model as it is without making any transformations and thereafter apply the perturbation technique of the previous section. It is not clear which estimator to prefer, i.e., the estimator obtained from reducing the number of variance components in the model, or the estimator obtained from the model consisting of all three variance components.

One solution to the problem is to compare the sampling variances of the two estimators. Thus, for model (12) we take two sets of estimation equations to estimate the variance components of the model. The first set of equations will be called Partition I, i.e., we estimate all three components $\sigma^2_1$, $\sigma^2_2$ and $\sigma^2_e$ as we have previously done in Section (2.3). In the second set of estimation equations we will have two variance components to estimate, $\sigma^2_1$ and $\sigma^2_e$, which will be called Partition II. The reduction in sum of squares which will be used to estimate the variance component for the first random effect $u_1$, denoted by $\widehat{\sigma}^2_1$, will be corrected for the fixed $\beta$ and second random effect $u_2$. Thereafter, we compare the variance of $\widehat{\sigma}^2_{u1}$ and the variance of $\widehat{\sigma}^2_1$ obtained from Partition I and Partition II, respectively.

## 4.1 Variance component estimator for Partition I

For model (12), we have already estimated the variance component estimators in section (2.3) with the set of estimation equations which were given by (20). For this, we needed the following matrices: $[X]$, $[X, Z_1]$ and $[X, Z_1, Z_2]$. Corresponding to these matrices, we had the projection matrices which were denoted by $P_x$, $P_{x1}$ and $P_{x_{12}}$, respectively. The variance component estimator of interest, $\widehat{\sigma}^2_{u1}$, was given in (23) and its variance, i.e., $\mathrm{D}[\widehat{\sigma}^2_{u1}]$, given in (27).

## 4.2 Variance component estimator for Partition II

For the same model (12), also using Henderson's method 3, we estimate $\sigma_1^2$, but this time with a different set of estimation equations. The set of equations which are needed, are the following SSE: The residual sum of squares. $\mathrm{R}(u_1/\beta, u_2)$: The reduction sums of squares due to the first random component, adjusted for the fixed and second random effect. To calculate $\mathrm{R}(u_1/\beta, u_2)$, we need the sub-model defined in (15). The corresponding projection matrix $P_{x2}$ is defined as,

$$P_{x_2} = (X, Z_2)((X, Z_2)'(X, Z_2))^-(X, Z_2)'. \tag{58}$$

We have

$$\mathrm{R}(u_1/\beta, u_2) = \mathrm{R}(\beta, u_1, u_2) - \mathrm{R}(\beta, u_2). \tag{59}$$

which gives

$$\begin{aligned}\mathrm{R}(u_1/\beta, u_2) &= Y'P_{x_{12}}Y - Y'P_{x_2}Y \\ &= Y'(P_{x_{12}} - P_{x_2})Y.\end{aligned}$$

i.e., the reduction in sum of squares needed are

$$\begin{cases}\mathrm{R}(u_1/\beta, u_2) \\ \mathrm{SSE}\end{cases}$$

Consequently, we can write the set of estimation equations to estimate the variance components $\sigma_1^2$ and $\sigma_e^2$ as,

$$\mathrm{E}\begin{bmatrix} Y'(P_{x_{12}} - P_{x_2})Y \\ Y'(I - P_{x_{12}})Y \end{bmatrix} = \begin{bmatrix} tr((P_{x_{12}} - P_{x_2})V_1) & tr(P_{x_{12}} - P_{x_2}) \\ tr((I - P_{x_{12}})V_1) & tr(I - P_{x_{12}}) \end{bmatrix}\begin{bmatrix} \sigma_1^2 \\ \sigma_e^2 \end{bmatrix}.$$

The right hand side of the matrix can be simplified, using $P_{x_{12}}V_1 = V_1$,

$$\begin{bmatrix} \widehat{\sigma}_1^2 \\ \widehat{\sigma}_e^2 \end{bmatrix} = J^{-1}\begin{bmatrix} Y'(P_{x_{12}} - P_{x_2})Y \\ Y'(I - P_{x_{12}})Y \end{bmatrix}, \tag{60}$$

where $J$ is equal to

$$\begin{bmatrix} tr(P_{x_{12}} - P_{x_2})V_1 & tr(P_{x_{12}} - P_{x_2}) \\ 0 & tr(I - P_{x_{12}}) \end{bmatrix}. \qquad (61)$$

Thus,

$$\begin{aligned} \widehat{\sigma}_1^2 &= \frac{\text{tr}(I - P_{x_{12}})Y'(P_{x_{12}} - P_{x_2})Y - \text{tr}(P_{x_{12}} - P_{x_2})Y'(I - P_{x_{12}})Y}{\text{tr}(P_{x_{12}} - P_{x_2})V_2 tr(I - P_{x_{12}})} \\ &= \frac{Y'(P_{x_{12}} - P_{x_2})Y}{\text{tr}(P_{x_{12}} - P_{x_2})V_1} - \frac{\text{tr}(P_{x_{12}} - P_{x_2})Y'(I - P_{x_{12}})Y}{\text{tr}(P_{x_{12}} - P_{x_2})V_1 \text{tr}(I - P_{x_{12}})}. \qquad (62) \end{aligned}$$

The variance is

$$\begin{aligned} D[\widehat{\sigma}_1^2] &= D[\frac{Y'(P_{x_{12}} - P_{x_2})Y}{tr(P_{x_{12}} - P_{x_2})V_1}] + D[\frac{tr(P_{x_{12}} - P_{x_2})Y'(I - P_{x_{12}})Y}{tr(P_{x_{12}} - P_{x_2})V_1 tr(I - P_{x_{12}})}] \\ &= [\frac{2tr(P_{x_{12}} - P_{x2})V_1(P_{x_{12}} - P_{x2})V_1}{(tr(P_{x_{12}} - P_{x2})V_1)^2}]\sigma_1^4 + \\ &\quad [\frac{4tr(P_{x_{12}} - P_{x2})^2 V_1}{(tr(P_{x_{12}} - P_{x2})V_1)^2}]\sigma_e^2\sigma_1^2 + [\frac{2tr(P_{x_{12}} - P_{x2})^2}{(tr(P_{x_{12}} - P_{x2})V_1)^2} + \\ &\quad \frac{2tr(P_{x_{12}} - P_{x2})^2}{(tr(P_{x_{12}} - P_{x2})V_1)^2 tr(I - P_{x_{12}})}]\sigma_e^4. \qquad (63) \end{aligned}$$

After calculating the variance component estimators, $\widehat{\sigma}_{u1}^2$ and $\widehat{\sigma}_1^2$ for Partition I and Partition II, respectively, we compare their variances. The estimator that has less variance can then be recommended and thereafter modified. Here it is essential to compare the leading terms of the coefficients of $\sigma_1^4$, $\sigma_e^4$ and $\sigma_2^4$. We are going to suppose that the variance component corresponding to the second random effect in Partition I, is considerably "small", because if this component is large, the variance function in (27) is going to be large as well. Under such circumstances, it could be better to take into consideration the estimators obtained from Partition II, instead of the ones obtained from Partition I.

The coefficient of $\sigma_1^4$ under Partition I equals

$$\frac{2}{(tr(P_{x1} - P_x)V_1)^2} tr(P_{x1} - P_x)V_1(P_{x1} - P_x)V_1,$$

and under Partition II

$$\frac{2}{(tr(P_{x_{12}} - P_{x2})V_1)^2} tr(P_{x_{12}} - P_{x2})V_1(P_{x_{12}} - P_{x2})V_1.$$

Hence, comparing the coefficients of $\sigma_1^4$ in the two partitions can be expressed in the following inequality

$$\frac{2}{(tr(P_{x1}-P_x)V_1)^2}tr(P_{x1}-P_x)V_1(P_{x1}-P_x)V_1$$
$$\leq \frac{2}{(tr(P_{x_{12}}-P_{x2})V_1)^2}tr(P_{x_{12}}-P_{x2})V_1(P_{x_{12}}-P_{x2})V_1. \quad (64)$$

Likewise, considering the coefficients of $\sigma_e^4$ in (27) and (63) we have the following

$$[\frac{2tr(P_{x1}-P_x)}{(\mathrm{tr}(P_{x1}-P_x)V_1)^2}+\frac{2(\mathrm{tr}(P_{x1}-P_x)V_2)^2\mathrm{tr}(P_{x_{12}}-P_{x1})^2}{(tr(P_{x1}-P_x)V_1)^2(tr(P_{x_{12}}-P_{x1})V_2)^2}$$
$$+\frac{2k^2}{(\mathrm{tr}(P_{x1}-P_x)V_1)^2(\mathrm{tr}(P_{x_{12}}-P_{x1})V_2)^2tr(I-P_{x_{12}})}],$$
$$\leq [\frac{2tr(P_{x_{12}}-P_{x2})^2}{(\mathrm{tr}(P_{x_{12}}-P_{x2})V_1)^2}+\frac{2(\mathrm{tr}(P_{x_{12}}-P_{x2}))^2}{(\mathrm{tr}(P_{x_{12}}-P_{x2})V_1)^2tr(I-P_{x_{12}})}]. \quad (65)$$

Thus,

**Proposition 2.** *In model (12) let the variance component corresponding to the first random effect be estimated according to the estimation equations given by (20) or (60), and denoted* $\widehat{\sigma}_{u1}^2$ *and* $\widehat{\sigma}_1^2$*, respectively. Then under the assumption that* $\sigma_1^2\sigma_e^2$ *and* $\sigma_2^2$ *are "small", (64) and (65) are sufficient conditions for* $D[\widehat{\sigma}_{u1}^2]\leq D[\widehat{\sigma}_1^2]$.

We can further simplify the expressions (64) and (65) as below: In (64), let $A=(P_{x1}-P_x)$ and $B=(P_{x_{12}}-P_{x2})$. Equation (64) can be written

$$\frac{\mathrm{tr}(AV_1AV_1)}{(trAV_1)^2}\leq\frac{\mathrm{tr}(BV_1BV_1)}{(\mathrm{tr}BV_1)^2}. \quad (66)$$

Since $V_1$ is symmetric, $V_1$ can be written $V_1=\Gamma D_1\Gamma'$ where $D_1$ is the diagonal matrix of the $r_1$ eigenvalues of $V_1$. Taking every part of (66) separately, the LHS can be written as

$$\begin{aligned}(\mathrm{tr}AV_1)^2 &= (\mathrm{tr}(A\Gamma D_1\Gamma'))^2\\ &= (\mathrm{tr}(A_\star D_1))^2\\ &= (\mathrm{tr}\begin{pmatrix}A_{\star_{11}} & A_{\star_{12}}\\ A_{\star_{21}} & A_{\star_{22}}\end{pmatrix}\begin{pmatrix}\Delta_1 & 0\\ 0 & 0\end{pmatrix})^2\\ &= (\mathrm{tr}(A_{\star_{11}}\Delta_1)^2\\ &= (\sum_{i=1}^{r_1}a_{\star ii}\lambda_{1i})^2, \quad (67)\end{aligned}$$

where $A_\star = \Gamma' A \Gamma$ and $\Delta_1 = \text{diag}(\lambda_{11}, \ldots, \lambda_{1r1})$, $r_1$ is the rank of $V_1$. For the term

$$\begin{aligned}
\text{tr}(AV_1AV_1) &= \text{tr}(A\Gamma D_1 \Gamma' A \Gamma D_1 \Gamma') \\
&= \text{tr}(A_\star D_1 A_\star D_1) \\
&= \text{tr}(\begin{pmatrix} A_{\star_{11}} & A_{\star_{12}} \\ A_{\star_{21}} & A_{\star_{22}} \end{pmatrix} \begin{pmatrix} \Delta_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_{\star_{11}} & A_{\star_{12}} \\ A_{\star_{21}} & A_{\star_{22}} \end{pmatrix} \begin{pmatrix} \Delta_1 & 0 \\ 0 & 0 \end{pmatrix}) \\
&= \text{tr}(A_{\star_{11}} \Delta_1 A_{\star_{11}} \Delta_1) = \sum_{i=1}^{r_1} a_{\star ii}^2 \lambda_{1i}^2 + 2 \sum_{i \neq j} a_{\star ij} a_{\star ji} \lambda_{1i} \lambda_{1j}. \qquad (68)
\end{aligned}$$

The right hand side of (66) can be written as

$$\begin{aligned}
(\text{tr}(BV_1))^2 &= (trB\Gamma D_1 \Gamma')^2 = (\text{tr} B_\star D_1)^2 = (\text{tr}(B_{\star_{11}} \Delta_1))^2 \\
&= (\sum_{i=1}^{r1} b_{\star ii} \lambda_{1i})^2, \qquad (69)
\end{aligned}$$

where $B_\star = \Gamma' B \Gamma$. Similar calculations give

$$\text{tr}(BV_1BV_1) = \sum_{i=1}^{r1} b_{\star ii}^2 \lambda_{1i}^2 + 2 \sum_{i \neq j} b_{\star ij} b_{\star ji} \lambda_{1i} \lambda_{1j}. \qquad (70)$$

Thus, (66) can now be rewritten as

$$\frac{\sum_{i=1}^{r_1} a_{\star ii}^2 \lambda_{1i}^2 + 2 \sum_{i \neq j} a_{\star ij} a_{\star ji} \lambda_{1i} \lambda_{1j}}{(\sum_{i=1}^{r_1} a_{\star ii} \lambda_{1i})^2} \leq \frac{\sum_{i=1}^{r1} b_{\star ii}^2 \lambda_{1i}^2 + 2 \sum_{i \neq j} b_{\star ij} b_{\star ji} \lambda_{1i} \lambda_{1j}}{(\sum_{i=1}^{r1} b_{\star ii} \lambda_{1i})^2}. \qquad (71)$$

Moreover, each part of (65) can be considered separately . First, from previous calculations the first term of the left hand side of (65) can be rewritten as

$$\frac{\text{tr}(P_{x1} - P_x)}{(\text{tr}(P_{x1} - P_x)V_1)^2} = \frac{r_{x1} - r_x}{(\sum_{i=1}^{r_1} a_{\star ii} \lambda_{1i})^2}. \qquad (72)$$

For the second part $V_2$ is involved, writing $V_2 = UD_2U'$, where $U$ is an orthogonal matrix, i.e., $UU' = I$ and $D_2$ is the diagonal matrix having the eigenvalues of $V_2$ on its diagonal. The rank of $V_2$ will be denoted by $r_2$

$$\begin{aligned}
\text{tr}((P_{x1} - P_x)V_2) &= \text{tr}(AV_2) = (\text{tr} AUD_2U') \\
&= (\text{tr} U'AUD_2) = \text{tr}(A_\bullet D_2) \\
&= \text{tr}(\begin{pmatrix} A_{\bullet_{11}} & A_{\bullet_{12}} \\ A_{\bullet_{21}} & A_{\bullet_{22}} \end{pmatrix} \begin{pmatrix} \Delta_2 & 0 \\ 0 & 0 \end{pmatrix}) = \sum_{i=1}^{r2} a_{\bullet ii} \lambda_{2i}.
\end{aligned}$$

where $A_\bullet = U'AU$ and $\Delta_2 = \mathrm{diag}(\lambda_{21}, \ldots, \lambda_{2r2})$. We also have the following term involved

$$(\mathrm{tr}(P_{x_{12}} - P_{x1})V_2)^2 = (\mathrm{tr}CV_2)^2 = (\mathrm{tr}C_\star D_2)^2 = (\sum_{i=1}^{r2} c_{\star ii}\lambda_{2i})^2, \tag{73}$$

where $C = (P_{x_{12}} - P_{x1})$ and $C_\star = U'CU$. The second term is equal to

$$\frac{((P_{x1} - P_x)V_2)^2 \mathrm{tr}(P_{x_{12}} - P_{x1})}{(\mathrm{tr}(P_{x1} - P_x)V_1)^2(\mathrm{tr}(P_{x_{12}} - P_{x1})V_2)^2} = \frac{(\sum_{i=1}^{r2} a_{\bullet ii}\lambda_{2i})^2(r_{x_{12}} - r_{x1})}{(\sum_{i=1}^{r1} a_{\star ii}\lambda_{1i})^2(\sum_{i=1}^{r2} c_{\star ii}\lambda_{2i})^2}. \tag{74}$$

Finally the third part is equal to

$$\frac{[\mathrm{tr}(P_{x1} - P_x)V_2\mathrm{tr}(P_{x_{12}} - P_{x1}) - \mathrm{tr}(P_{x_1} - P_x)\mathrm{tr}(P_{x_{12}} - P_{x1})V_2]^2}{(\mathrm{tr}(P_{x1} - P_x)V_1)^2(\mathrm{tr}(P_{x_{12}} - P_{x1})V_2)^2\mathrm{tr}(I - P_{x_{12}})}$$

$$= \frac{[\mathrm{tr}AV_2(r_{x_{12}} - r_{x1}) - (r_{x_1} - r_x)\mathrm{tr}CV_2]^2}{(\mathrm{tr}AV_1)^2(\mathrm{tr}CV_2)^2(n - r_{x_{12}})}$$

$$= \frac{[(r_{x_{12}} - r_{x1})\sum_{i=1}^{r2} a_{\bullet ii}\lambda_{2i} - (r_{x_1} - r_x)\sum_{i=1}^{r2} c_{\star ii}\lambda_{2i}]^2}{(\sum_{i=1}^{r1} a_{\star ii}\lambda_{1i})^2(\sum_{i=1}^{r2} c_{\star ii}\lambda_{2i})^2(n - r_{x_{12}})}. \tag{75}$$

Performing the same calculations as above, for the right hand side of (65) we get the following

$$\frac{\mathrm{tr}(P_{x_{12}} - P_{x2})}{(\mathrm{tr}(P_{x_{12}} - P_{x2})V_1)^2} + \frac{(\mathrm{tr}(P_{x_{12}} - P_{x2}))^2}{(\mathrm{tr}(P_{x_{12}} - P_{x2})V_1)^2 tr(I - P_{x_{12}})}$$

$$= \frac{(r_{x_{12}} - r_{x2})}{(\sum_{i=1}^{r1} b_{\star ii}\lambda_{1i})^2} + \frac{(r_{x_{12}} - r_{x2})^2}{(\sum_{i=1}^{r1} b_{\star ii}\lambda_{1i})^2(n - r_{x_{12}})}. \tag{76}$$

We have now all the involved terms for the coefficient of $\sigma_e^4$ for the two partitions, i.e., (27) and (63). Thus the (65) can be written as

$$\Bigg[\frac{r_{x1} - r_x}{(\sum_{i=1}^{r_1} a_{\star ii}\lambda_{1i})^2} + \frac{(\sum_{i=1}^{r2} a_{\bullet ii}\lambda_{2i})^2(r_{x_{12}} - r_{x1})}{(\sum_{i=1}^{r1} a_{\star ii}\lambda_{1i})^2(\sum_{i=1}^{r2} c_{\star ii}\lambda_{2i})^2}$$

$$+ \frac{[(r_{x_{12}} - r_{x1})\sum_{i=1}^{r2} a_{\bullet ii}\lambda_{2i} - (r_{x_1} - r_x)\sum_{i=1}^{r2} c_{\star ii}\lambda_{2i}]^2}{(\sum_{i=1}^{r1} a_{\star ii}\lambda_{1i})^2(\sum_{i=1}^{r2} c_{\star ii}\lambda_{2i})^2(n - r_{x_{12}})}\Bigg]$$

$$\leq \Bigg[\frac{(r_{x_{12}} - r_{x2})}{(\sum_{i=1}^{r1} b_{\star ii}\lambda_{1i})} + \frac{(r_{x_{12}} - r_{x2})^2}{(\sum_{i=1}^{r1} b_{\star ii}\lambda_{1i})^2(n - r_{x_{12}})}\Bigg] \tag{77}$$

To summarize:
Comparing the variances, $D[\widehat{\sigma}_{u1}^2]$ and $D[\widehat{\sigma}_1^2]$, obtained from the two different estimation equations denoted by Partition I and II, we have found certain conditions under which any of the estimators can be preferred, i.e., the estimator that has less variance.

To examine further (64) and (65), different examples will be considered. We have studied a few examples and calculated the value of the inequalities (64) and (65) corresponding to the variance functions, i.e., (27) and (63). The considered examples have different numbers of observations $n$ and the data have different experimental design patterns. From Table 1, we can observe according to the calculated values of the inequalities when data is balanced as in Examples 1, 2 and 8, that there seems to be no difference as to which partition to apply as was expected. In all the other examples, both (64) and (65) were satisfied indicating that Partition II should be recommended.

## 5 Conclusion

The problem of modifying the variance component estimator obtained by using Henderson's method 3, has been the focus of our work as well as to compare two different decompositions of sums of squares.

For a two-way linear mixed model, consisting of three variance components, $\sigma_1^2$, $\sigma_2^2$ and $\sigma_e^2$, we have perturbed the Henderson's estimation equations. The main aim, was to modify the standard unbiased estimator, corresponding to one of the random effects, by multiplying the estimator with some coefficients that are chosen to minimize the leading terms, $\sigma_1^4$, $\sigma_2^4$ and $\sigma_e^4$ in the MSE equation. Two modified variance component estimators are proposed; each appropriate under certain given conditions. Our proposed estimators are easy to compute and have smaller MSE than the unmodified one. Moreover, the conditions under which each of the proposed estimators are valid, are easy to investigate. For instance, in practical application if the unbiasedness condition is not of major concern, our proposed estimators should be considered.

We have studied two decompositions of Henderson's method 3, which we denoted by Partition I and Partition II. The former consisting of three variance components, and the latter of two variance components. The variances of the variance component estimators $\widehat{\sigma}_{u1}^2$ and $\widehat{\sigma}_1^2$ obtained from the two partitions, were compared under the assumption that $\sigma_2^2$ and $\sigma_1^2\sigma_e^2$ are "small".

Examples consisting of balanced and unbalanced data are considered. With balanced data, the partitions are equal . Otherwise, when data are unbalanced we have two cases; for the case when Partition I is recommended our modification approach is suitable. For cases when Partition II can be recommended, we refer to the paper by Kelly and Mathew (1994).

## References


[1] Graybill, F.A. and Hultquist, R.A. (1961). Theorems concerning Eisenhart's Model II. *Ann. Math. Stat.* **32**, 261-269.

[2] Henderson, C.R. (1953). Estimation of variance and covariance components. *Biometrics* **9**, 226-252.

[3] Kelly, J.R. and Mathew, T. (1994). Improved nonnegative estimation of variance components in some mixed models with unbalanced data. *Technometrics*, **36**, 171-181.

[4] Kuri, A.I., Mathew, T. and Sinha, B.K. (1998). *Statistical Tests For Mixed Linear Models.* John Wiley & Sons, New York.

[5] LaMotte, L.R. (1973). On non-negative quadratic unbiased estimation of variance components. *J. Amer. Stat., Assoc.,* **68**, 728-730.

[6] Schott, J.R. (1997). *Matrix Analysis For Statistics.* John Wiley & Sons, New York.

[7] Searle, S.R. (1968). Another look at Henderson's methods of estimating variance components. *Biometrics*, **24,** 749-778.

[8] Searle, S.R. (1971). *Linear Models*, John Wiley & Sons, New York.

[9] Searle, S.R. (1987). *Linear Models For Unbalanced Data.* John Wiley & Sons, New York.

[10] Searle, S.R. Casella, G., and McCulloch, C.E. (1992). *Variance Components.* John Wiley & Sons, New York.

Table 1: Different examples, coefficients of the leading terms $\sigma_1^4$ and $\sigma_e^4$ are compared for Partition I and Partition II

| Example | The Model | $n$ | eq:(64) | eq:(65) |
|---|---|---|---|---|
| 1 | $Y = 1_8\mu + \begin{pmatrix} 1_4 & 0 \\ 0 & 1_4 \end{pmatrix} u_1 + \begin{pmatrix} 1_2 & 0 \\ 0 & 1_2 \\ 1_2 & 0 \\ 0 & 1_2 \end{pmatrix} u_2 + e.$ | 8 | equal | equal |
| 2 | $Y = 1_{30}\mu + \begin{pmatrix} 1_{15} & 0 \\ 0 & 1_{15} \end{pmatrix} u_1 + \begin{pmatrix} 1_5 & 0 & 0 \\ 0 & 1_5 & 0 \\ 0 & 0 & 1_5 \\ 1_5 & 0 & 0 \\ 0 & 1_5 & 0 \\ 0 & 0 & 1_5 \end{pmatrix} u_2 + e.$ | 30 | equal | equal |
| 3 | $Y = 1_8\mu + \begin{pmatrix} 1_5 & 0 \\ 0 & 1_3 \end{pmatrix} u_1 + \begin{pmatrix} 1_2 & 0 \\ 0 & 1_3 \\ 1_1 & 0 \\ 0 & 1_2 \end{pmatrix} u_2 + e.$ | 8 | equal | satisfied |
| 4 | $Y = 1_8\mu + \begin{pmatrix} 1_6 & 0 \\ 0 & 1_2 \end{pmatrix} u_1 + \begin{pmatrix} 1_4 & 0 \\ 0 & 1_2 \\ 1_1 & 0 \\ 0 & 1_1 \end{pmatrix} u_2 + e.$ | 8 | equal | satisfied |
| 5 | $Y = 1_{30}\mu + \begin{pmatrix} 1_{10} & 0 & 0 \\ 0 & 1_{15} & 0 \\ 0 & 0 & 1_5 \end{pmatrix} u_1 + \begin{pmatrix} 1_5 & 0 & 0 \\ 0 & 1_5 & 0 \\ 1_{10} & 0 & 0 \\ 0 & 1_5 & 0 \\ 0 & 1_2 & 0 \\ 0 & 0 & 1_3 \end{pmatrix} u_2 + e.$ | 30 | satisfied | satisfied |
| 6 | $Y = 1_{30}\mu + \begin{pmatrix} 1_7 & 0 & 0 & 0 \\ 0 & 1_{12} & 0 & 0 \\ 0 & 0 & 1_6 & 0 \\ 0 & 0 & 0 & 1_5 \end{pmatrix} u_1 + \begin{pmatrix} 1_4 & 0 & 0 \\ 0 & 0 & 1_3 \\ 0 & 1_{10} & 0 \\ 0 & 0 & 1_2 \\ 1_2 & 0 & 0 \\ 0 & 1_4 & 0 \\ 1_5 & 0 & 0 \end{pmatrix} u_2 + e.$ | 30 | satisfied | satisfied |
| 7 | $Y = 1_{21}\mu + \begin{pmatrix} 1_5 & 0 & 0 \\ 0 & 1_9 & 0 \\ 0 & 0 & 1_7 \end{pmatrix} u_1 + \begin{pmatrix} 1_2 & 0 & 0 \\ 0 & 1_3 & 0 \\ 0 & 1_1 & 0 \\ 0 & 0 & 1_8 \\ 1_4 & 0 & 0 \\ 0 & 1_3 & 0 \end{pmatrix} u_2 + e$ | 21 | satisfied | satisfied |
| 8 | $Y = 1_{30}\mu + \begin{pmatrix} 1_{10} & 0 & 0 \\ 0 & 1_{10} & 0 \\ 0 & 0 & 1_{10} \end{pmatrix} u_1 + \begin{pmatrix} 1_5 & 0 \\ 0 & 1_5 \\ 1_5 & 0 \\ 0 & 1_5 \\ 1_5 & 0 \\ 0 & 1_5 \end{pmatrix} u_2 + e$ | 30 | equal | equal |